Markov chain approach for computing polarized radiative transfer (RT):

formalism and application

Feng Xu

Anthony B. Davis

Jet Propulsion Laboratory

California Institute of Technology

Jet Propulsion Laboratory
California Institute of Technology

23 Mar 2017

Markov chain RT modeling: history

THE ASTROPHYSICAL JOURNAL, 219: 1058–1067, 1978 February 1

1 st MarCh paper

RADIATIVE TRANSFER CALCULATED FROM A MARKOV CHAIN FORMALISM*

LARRY W. ESPOSITO
University of Massachusetts, Amherst
AND
LEWIS L. HOUSE
High Altitude Observatory, National Center for Atmospheric Research†
Received 1977 April 18; accepted 1977 August 5

ABSTRACT

The theory of Markov chains is used to formulate the radiative transport problem in a general way by modeling the successive interactions of a photon as a stochastic process. Under the minimal requirement that the stochastic process is a Markov chain, the determination of the diffuse reflection or transmission from a scattering atmosphere is equivalent to the solution of a system of linear equations. This treatment is mathematically equivalent to, and thus has many of the advantages of, Monte Carlo methods, but can be considerably more rapid than Monte Carlo algorithms for numerical calculations in particular applications. We have verified the speed and accuracy of this formalism for the standard problem of finding the intensity of scattered light from a homogeneous plane-parallel atmosphere with an arbitrary phase function for scattering. Accurate results over a wide range of parameters were obtained with computation times comparable to those of a standard "doubling" routine. The generality of this formalism thus allows fast, direct solutions to problems that were previously soluble only by Monte Carlo methods. Some comparisons are made with respect to integral equation methods.

Subject headings: planets: atmospheres — radiative transfer

Method	Rayleigh Atmosphere (τ=84)	Inhomogeneous Venus Atmosphere
Adding or doubling	0.8 seconds	370 seconds
Markov chain	1.2 seconds	18 seconds

Larry W. Esposito

Lab for Atmos. & Space Phys.

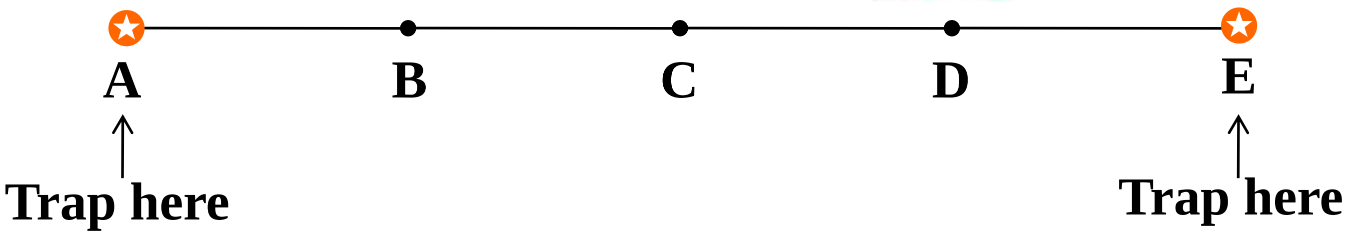
University of Colorado

Venus

Algorithm concept

Drunk's walk – a toy model:

- **Block**

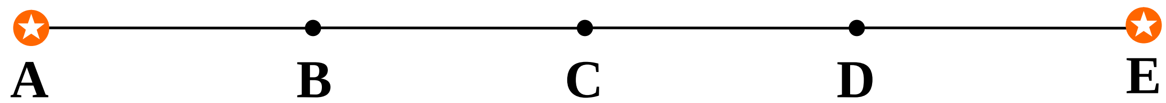


Starting from D, what is probability for him to get trapped at A and E, respectively ?

Probability matrix:

	A	B	C	D	E
A	1	0	0	0	0
B	0.5	0	0.5	0	0
C	0	0.5	0	0.5	0
D	0	0	0.5	0	0.5
E	0	0	0	0	1

Algorithm concept (cont'd)



Step 1: Matrix re-arrangement:

$$
\begin{array}{c|ccc|cc}
 & B & C & D & A & E \\
\hline
B & 0 & 0.5 & 0 & 0.5 & 0 \\
C & 0.5 & 0 & 0.5 & 0 & 0 \\
D & 0 & 0.5 & 0 & 0 & 0.5 \\
\hline
A & 0 & 0 & 0 & 1 & 0 \\
E & 0 & 0 & 0 & 0 & 1
\end{array}
$$

Q R

Q: transition matrix;

R: absorbing matrix;

"Intermediate" status: Block B, C and D;

"Absorbing" status: Block A and E.

Step 2: Solution to the probability of getting trapped at A or E:

$$\mathbf{X=R(E-Q)^{(-1)}\,I_0}$$

Pr(A) Pr(E)

E: identity matrix;

$\mathbf{I}_0=[1,0,0]^T$ for the drunk is at Block B; X = $[75\%, 25\%]^T$

$\mathbf{I}_0=[0,1,0]^T$ for the drunk is at Block C; X = $[50\%, 50\%]^T$

$\mathbf{I}_0=[0,0,1]^T$ for the drunk is at Block D. X = $[25\%, 75\%]^T$

X(1): probability of being trapped at A;

X(2): probability of being trapped at E.

$$\mathbf{(E-Q)^{(-1)}=E+Q+QQ+QQQ+\ldots}$$

$$\mathbf{X=RI_0+RQI_0+RQQI_0+RQQQI_0+\ldots}$$

Transition to atmospheric scattering

$$I = R(E-Q)^{(-1)}I_0$$

$$= RI_0 + RQI_0 + RQQI_0 + \ldots$$

1: Status of photon: (n, i)

n: the layer number where photon stays

i: the direction it is going (θ_i)

2: Transition matrix $Q_{(n',j')|(n,i)}$

Transition probability of a photon from one intermediate status (n, i) to another (n', j);

3: Absorbing matrix $R_{(e)|(n',j)}$

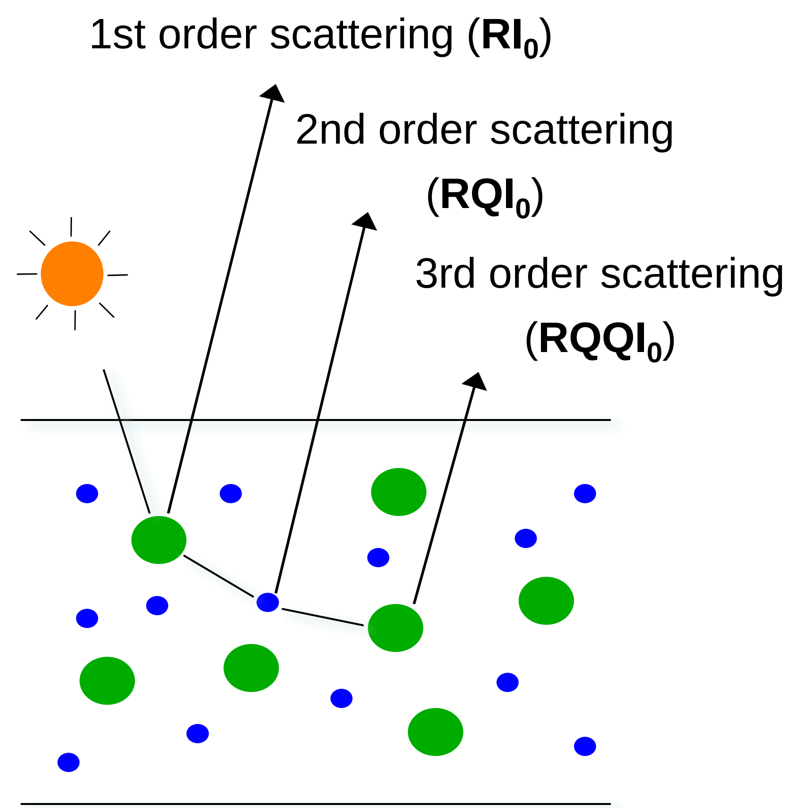
Probability of a photon escaping from status (n', j) to be out of the atmosphere in direction (θ_e);

4: Initial probability distribution I_0

Photon's initial distribution in status (n_0, i_0)

5: Solution vector I

The photon's probability to be emergent in direction (θ_e)

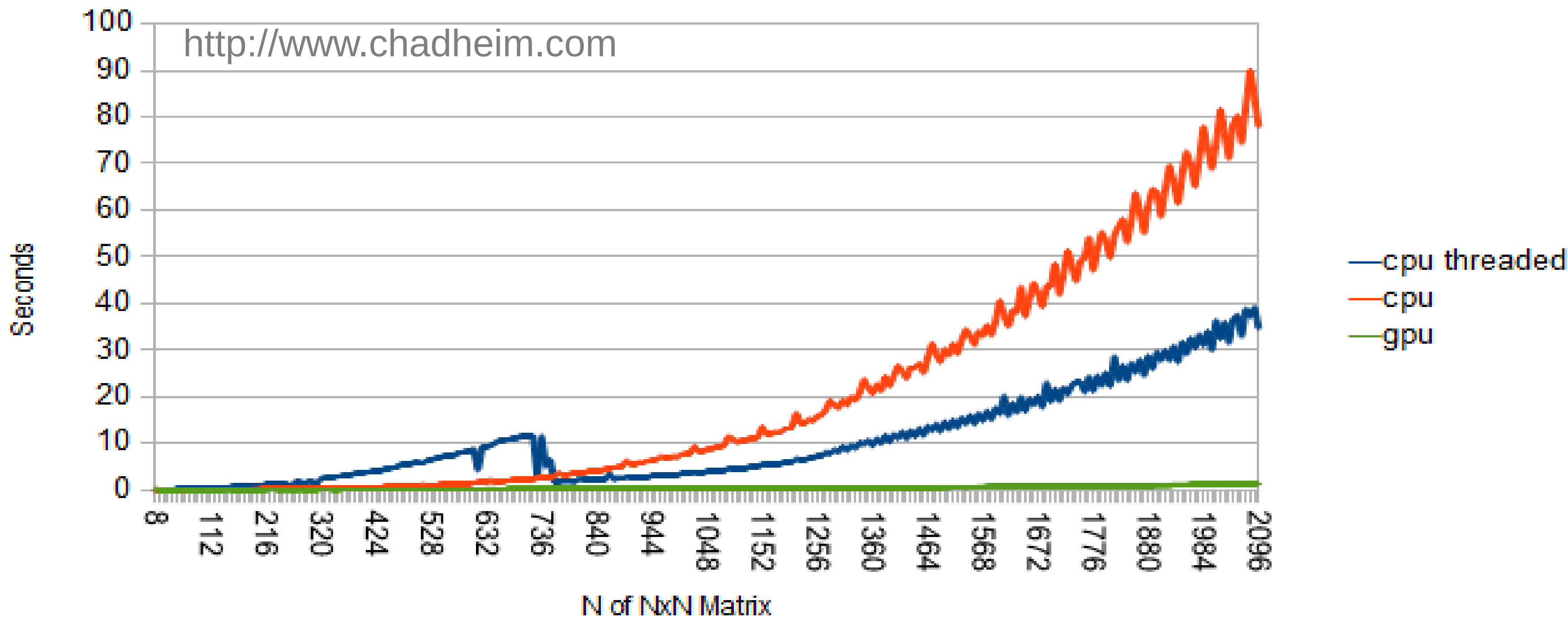


Characteristics of matrix algebra

- ✓ The physical sense of each matrix multiplication series is clear
- ✓ Matrix inverse computation can be recycled for Jacobian computation

$$\frac{\partial \mathbf{I}}{\partial a} = \frac{\partial \mathbf{R}}{\partial a}(\mathbf{E}-\mathbf{Q})^{-1} + \mathbf{R}(\mathbf{E}-\mathbf{Q})^{-1}\frac{\partial \mathbf{Q}}{\partial a}(\mathbf{E}-\mathbf{Q})^{-1}$$

- ✓ The matrix multiplication basis lends Markov chain suitable to implement on GPU

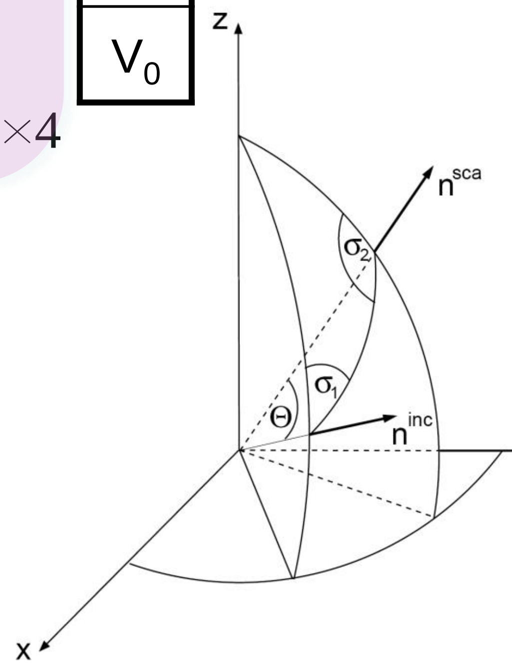


Extension to polarized RT

Scalar RT: $I = P_{11} \times I_0$

Vector RT:

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \left(R(\sigma_2) \begin{bmatrix} P_{11} & P_{21} & 0 & 0 \\ P_{21} & P_{22} & 0 & 0 \\ 0 & 0 & P_{33} & -P_{43} \\ 0 & 0 & P_{43} & P_{44} \end{bmatrix} R(\sigma_1) \right)_{4\times 4} \begin{bmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{bmatrix}$$



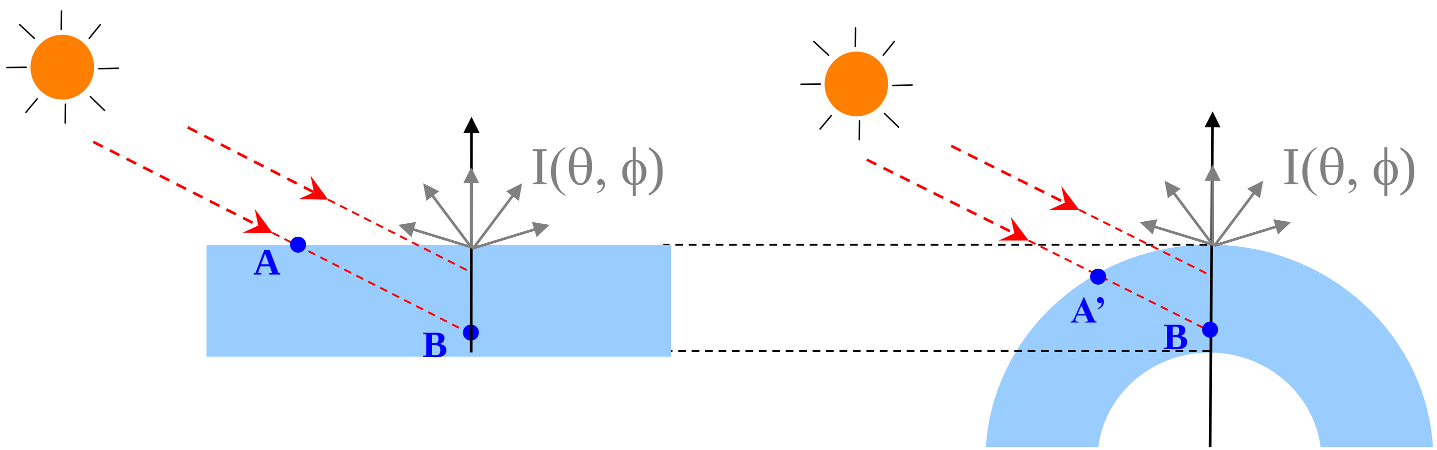
Surface reflection model

Total reflection = BRDF(λ, μ_{Sun}, μ_{View}, $\Delta\phi$) + pBRDF(μ_{Sun}, μ_{View}, $\Delta\phi$)

$$r_{11,depo} = a_\lambda \text{ x } r_{depo}(\mu_{Sun}, \mu_{View}, \Delta\phi) \qquad \mathbf{r}_{po} = \varepsilon \times f(\mathbf{P}_{po})$$

Extension to spherical-shell atmosphere

- Develop **Markov chain method** for computing polarized RT in plane-parallel (P-P) atmosphere
- Pseudo-sphericalization of the P-P RT model for accounting atmosphere sphericity in approximate manner
- Combined with Picard iteration for accurate RT field

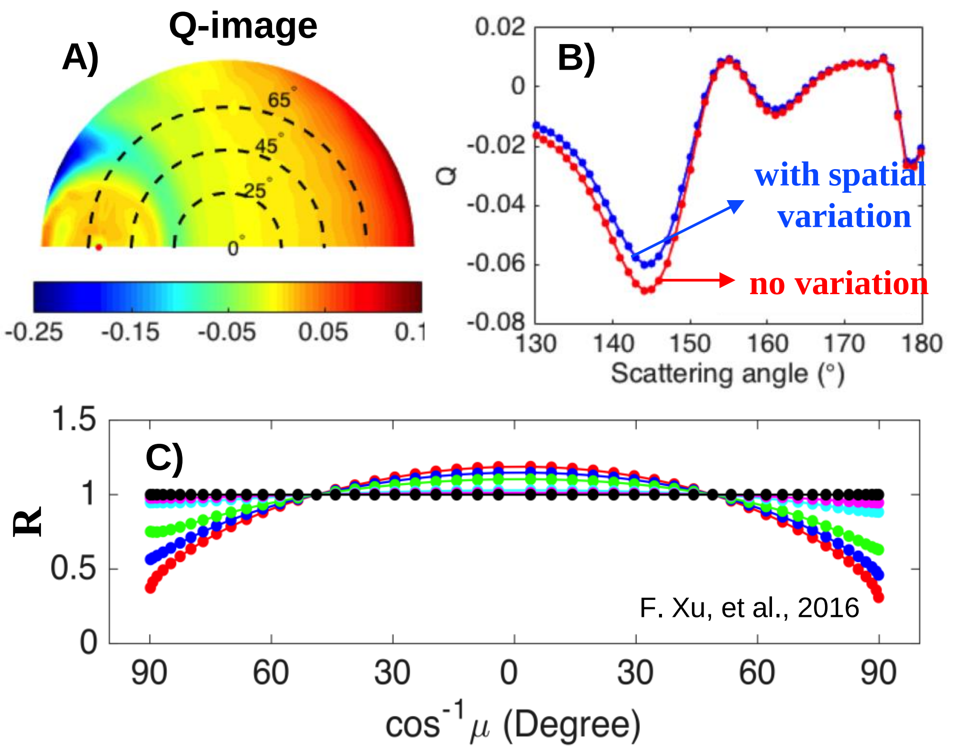


Incident attenuation integral path: **A to B**

Incident attenuation integral path: **A' to B**

Xu et al, JQSRT (2013).

Dealing with random particle density fluctuation



A): Simulated angular distribution of Stokes vector Q (radial: viewing angle) in the existence of spatial variation; B): Angular distribution of polarized cloudbow signal with and wrt spatial variation; C): Non-unity BRF ratio (R vs angles, colors for different levels of spatial heterogeneity of cloud droplet density) shows the violation of reciprocity.

Problem: Spatial heterogeneity is the reality of planetary atmosphere. Accurate cloud remote sensing in this situation relies on a fast yet accurate radiative transfer (RT) solver to the relevant generalized RT equation (gRTE, Davis and Xu, 2014).

Finding: By developing a Markov chain solution to the gRTE, we find (i) angular reciprocity with classical RT models is violated to a degree that increases with spatial variability; (ii) angular positions of cloudbows, supernumerary bows, and glories are relatively unchanged.

Significance: The computational efficiency and accuracy of Markov chain solution has the potential to enable an efficient aerosol and cloud droplet retrieval scheme in an atmospheric medium with spatial heterogeneity.

Linearization for optimization

Forward model: $\mathbf{I} = \mathbf{R}\,(\mathbf{E} - \mathbf{Q})^{-1} \times \mathbf{I}_s$

Linearization: $$\frac{\partial \mathbf{I}}{\partial a} = \frac{\partial \mathbf{R}}{\partial a}(\mathbf{E}-\mathbf{Q})^{-1} + \mathbf{R}(\mathbf{E}-\mathbf{Q})^{-1}\frac{\partial \mathbf{Q}}{\partial a}(\mathbf{E}-\mathbf{Q})^{-1}$$

Jacobian matrix:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{I}_1}{\partial a_1} & \frac{\partial \mathbf{I}_1}{\partial a_2} & \cdots & \frac{\partial \mathbf{I}_1}{\partial a_N} \\ \frac{\partial \mathbf{I}_2}{\partial a_1} & \frac{\partial \mathbf{I}_2}{\partial a_2} & \cdots & \frac{\partial \mathbf{I}_2}{\partial a_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial \mathbf{I}_M}{\partial a_1} & \frac{\partial \mathbf{I}_M}{\partial a_2} & \cdots & \frac{\partial \mathbf{I}_M}{\partial a_N} \end{bmatrix}_{M \times N}$$

$\mathbf{I}_{obs}$: observations

$\mathbf{W}$: weighting matrix

$\Delta\mathbf{a}$: increment of solution vector

$\mathbf{a}$: solution vector

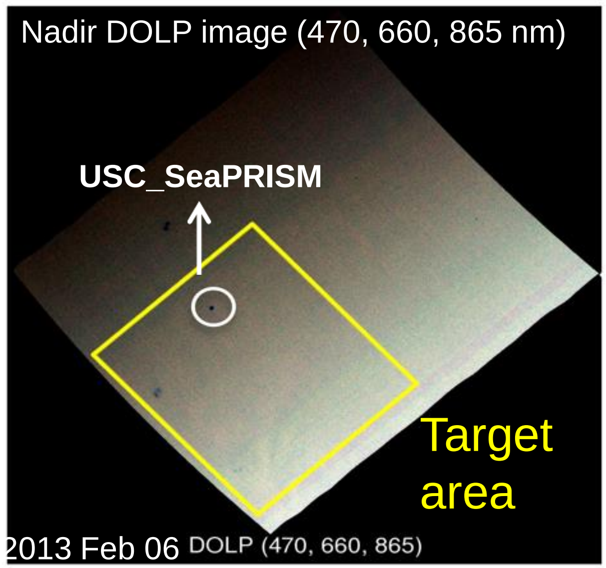
λ: damping factor

Levenberg algorithm: $$\Delta\mathbf{a}_k = \left[\mathbf{J}_k^{\mathrm{T}}\mathbf{W}\mathbf{J}_k + \lambda_k \operatorname{diag}\left(\mathbf{J}_k^{\mathrm{T}}\mathbf{W}\mathbf{J}_k\right)\right]^{-1}\left[\mathbf{J}_k^{\mathrm{T}}\mathbf{W}(\mathbf{I}_k - \mathbf{I}_{obs})\right]$$

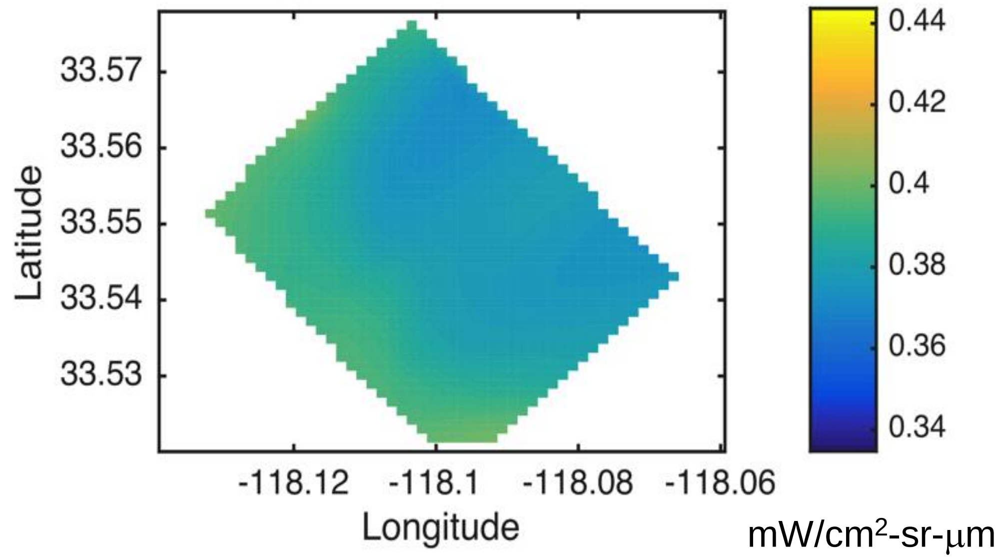
Xu et al., Appl. Opt. (2012).

Applications

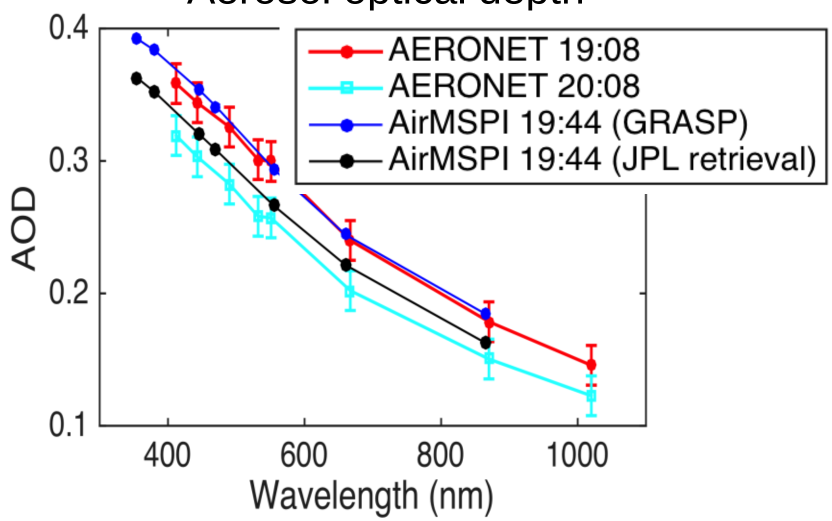
Coupled water-leaving radiance and aerosol retrieval



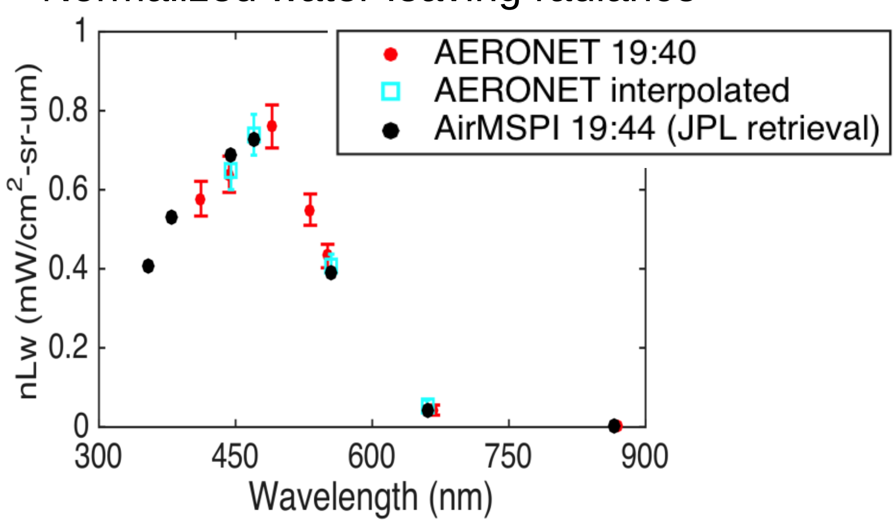
Normalized water-leaving radiance (555 nm)



Aerosol optical depth

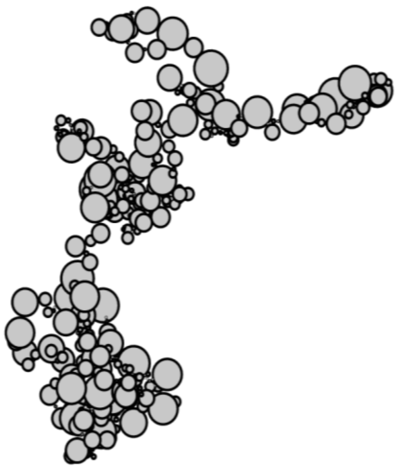


Normalized water-leaving radiance

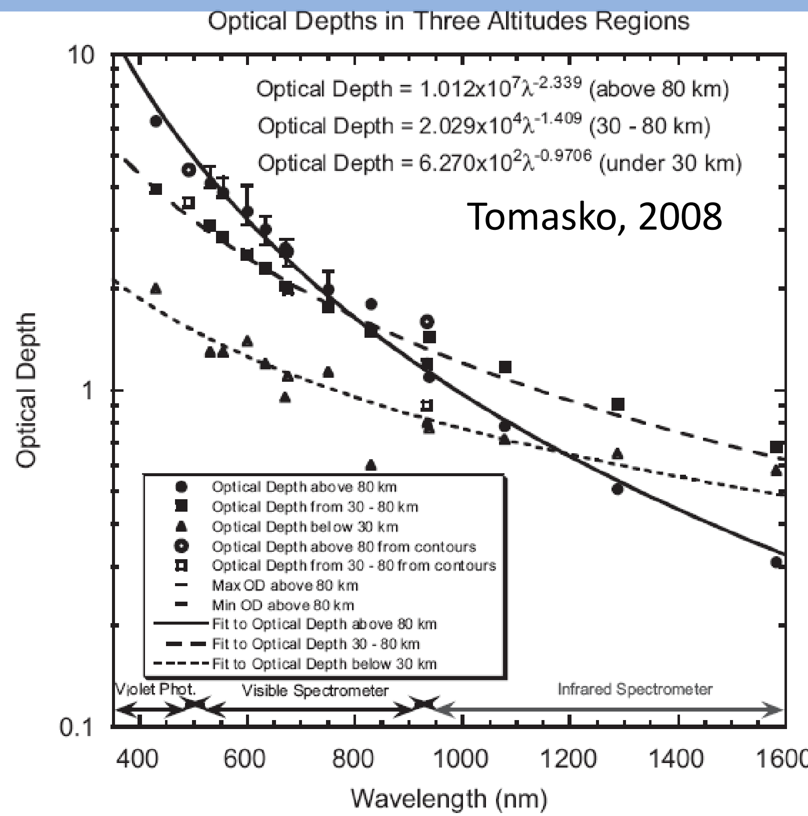


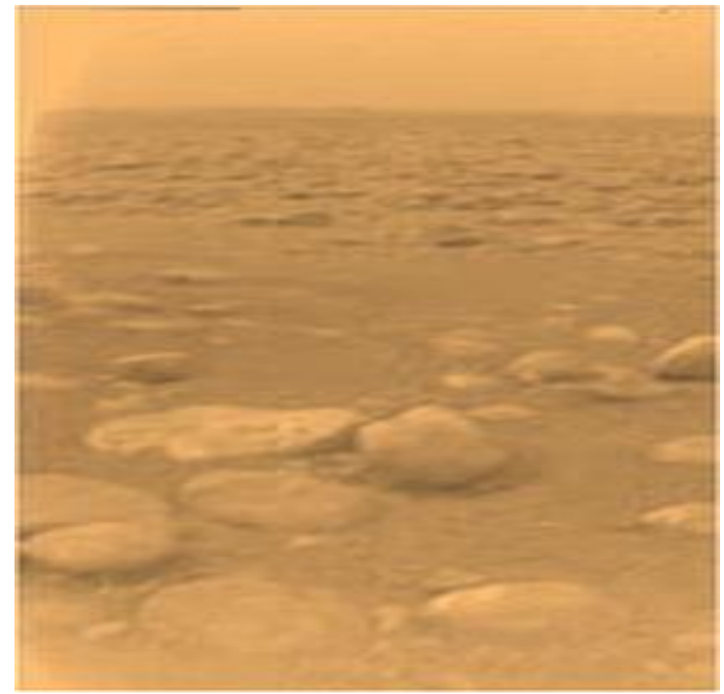
Titan haze aerosols and surface

West et al. 1991



Titan's aerosol (Tholin)



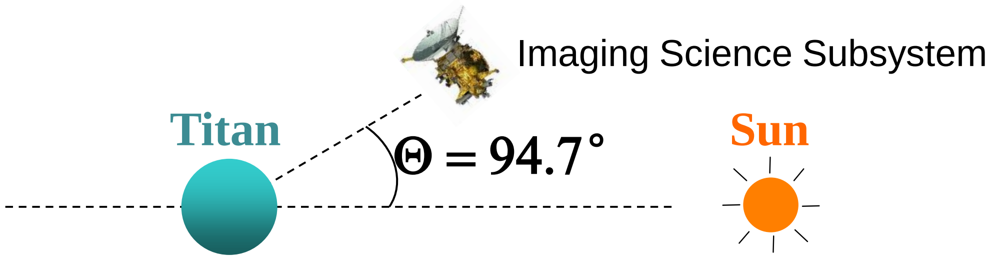


Schroder & Keller (2009) used Hapke's model (1993) to quantify Titan's surface reflection

$$r_F(\mu_0,\mu,\alpha)=\frac{w}{4}\frac{\mu_0}{\mu_0+\mu}\left[P(\alpha)B_{SH}(\alpha)+M(\mu_0,\mu)\right]B_{CB}(\alpha)S_{MR}(\mu_0,\mu,\varphi,\Theta)$$

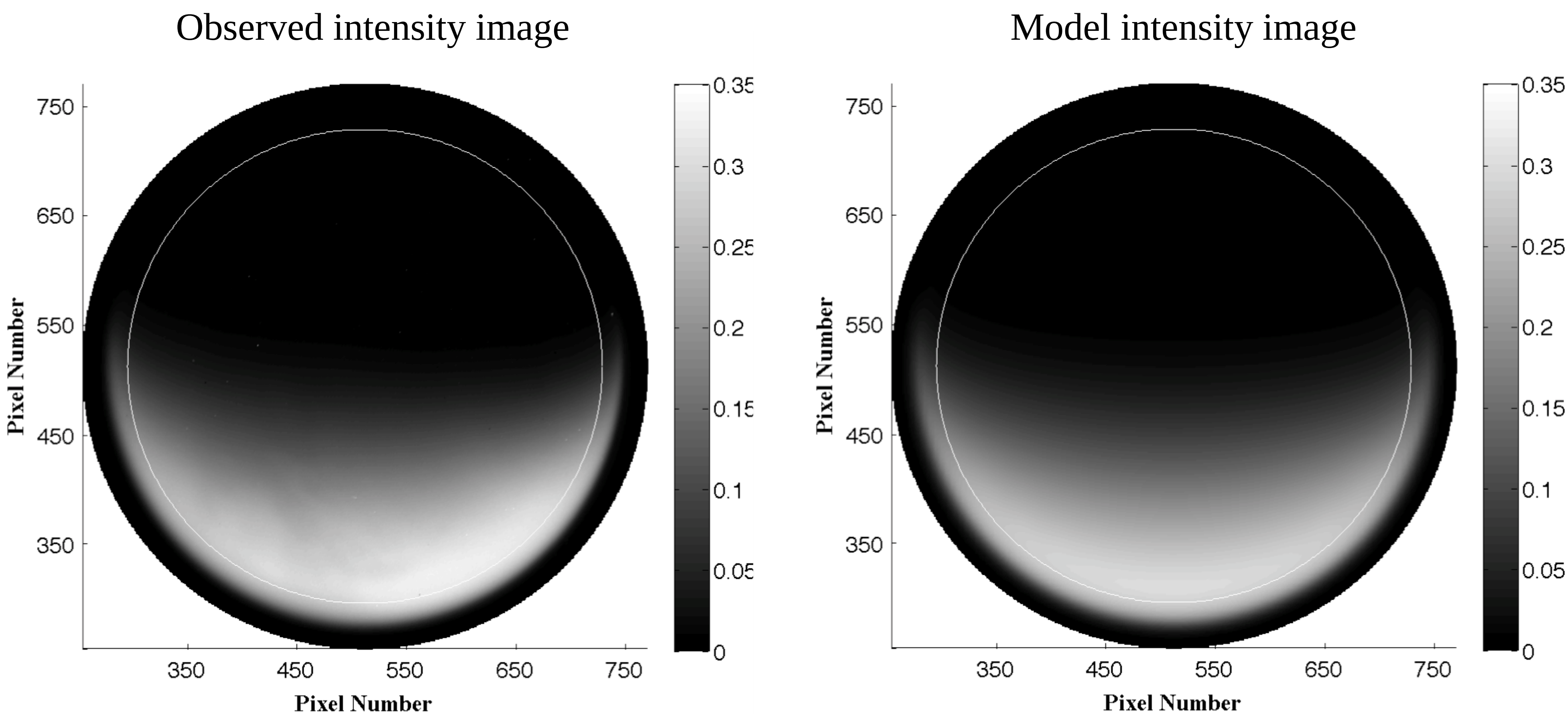
SH: shadowing function; CB: coherent backscattering; MR: macroscopic roughness

Model Titan's intensity image

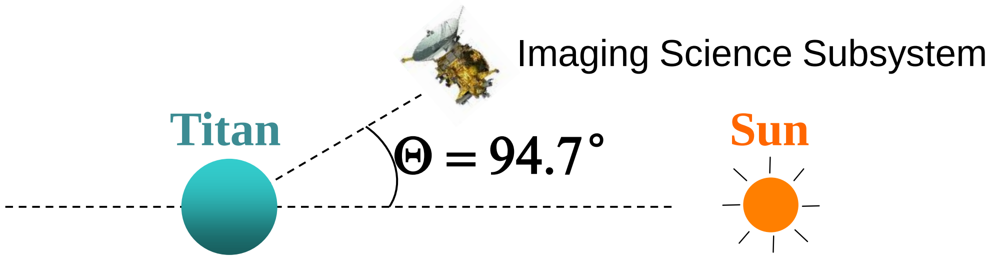


Titan on day 309 of 2008

Wavelength: 934.8 nm

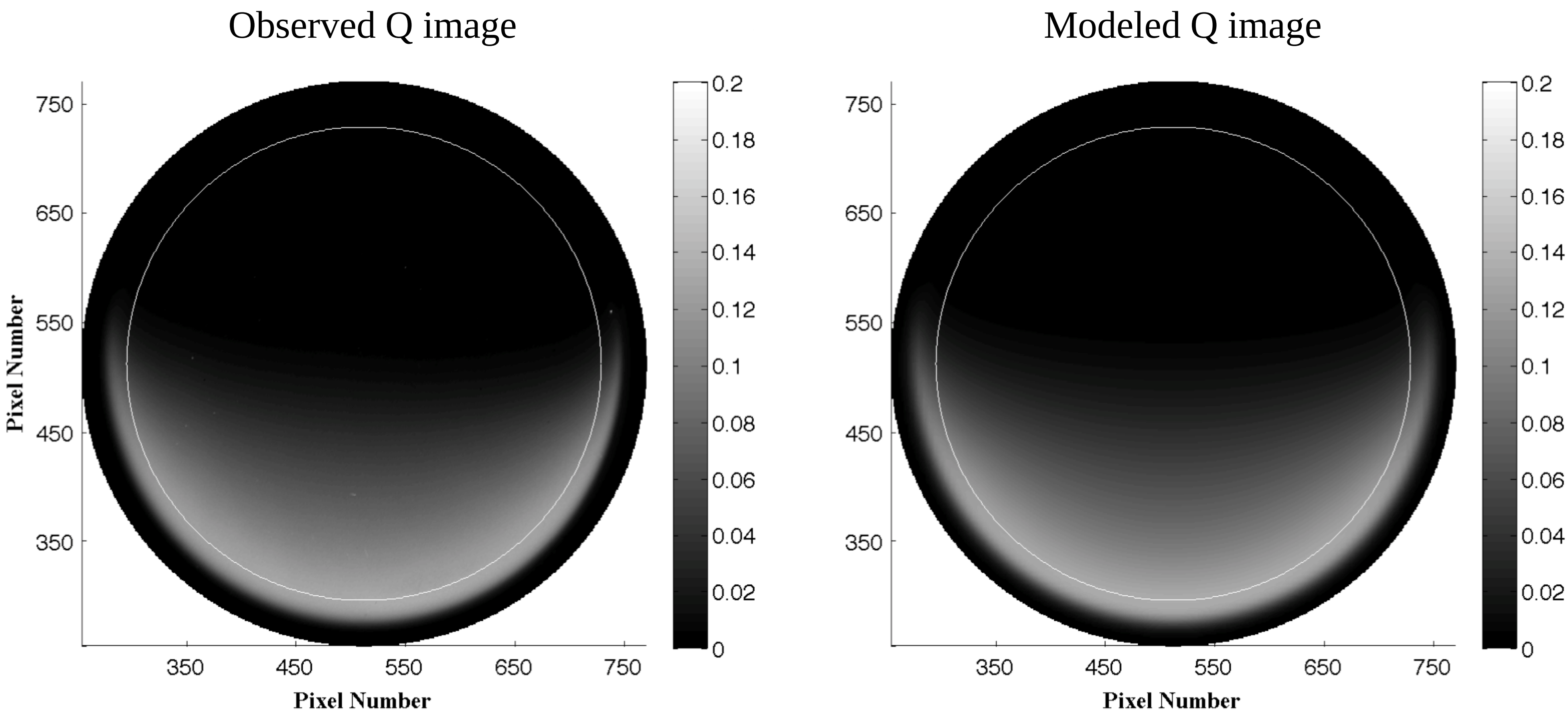


Model Titan's Q-image

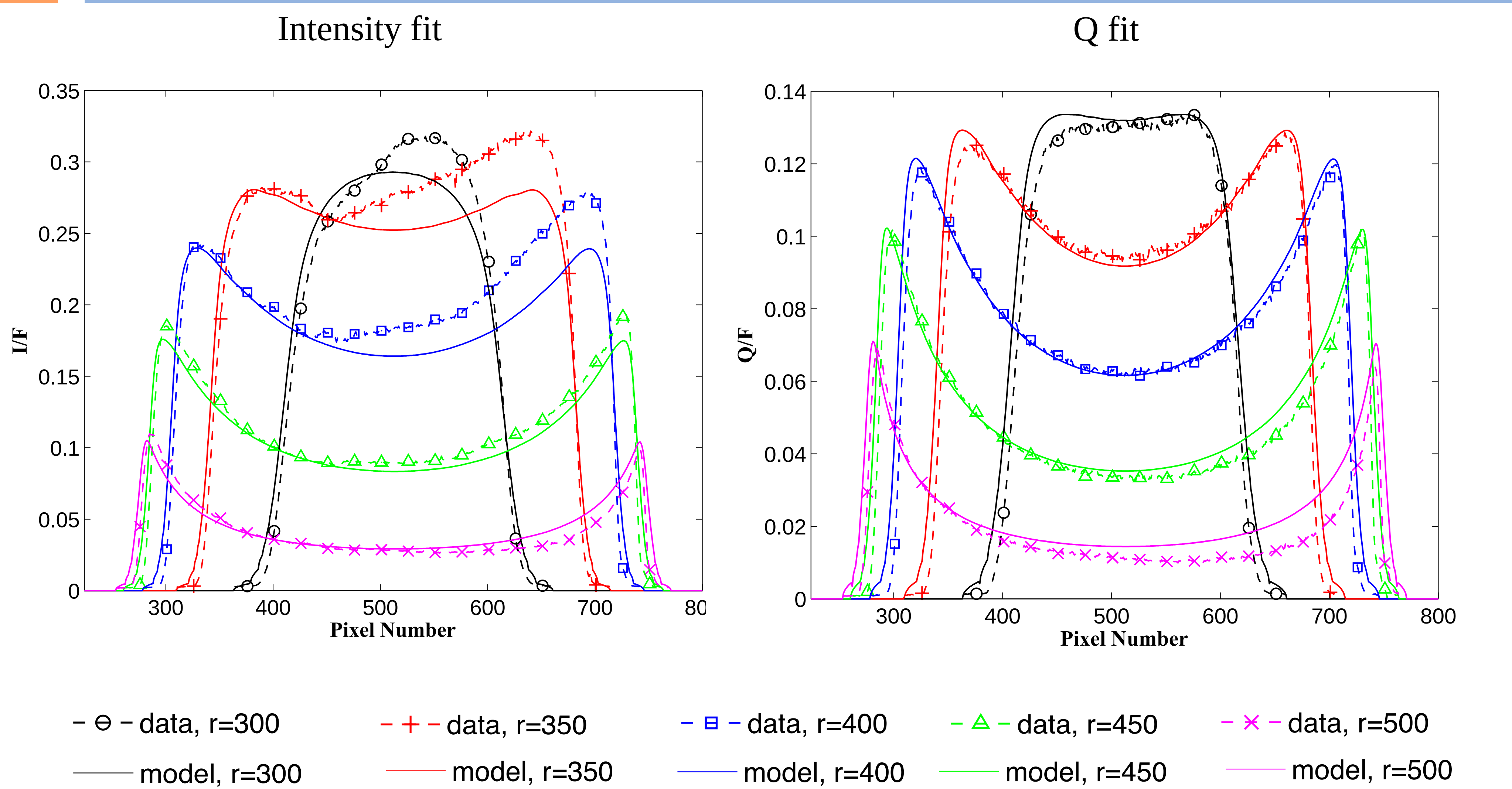


Titan on day 309 of 2008

Wavelength: 934.8 nm



Detailed comparison

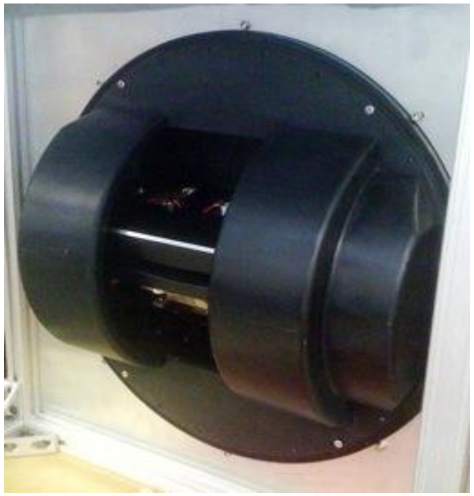


Major progress summary

When	What is done ?	By whom ?
1978	Unpolarized RT in homogeneous atmosphere	Esposito, House
1979	Unpolarized RT in plane-parallel atmosphere	Esposito
2010	Polarized Markov chain model	Xu, Esposito, West
2011	Various surface reflection model included	Xu, West
2012	Linearization complete	Xu, Davis
2012	Spherical-shell atmosphere (combined with Picard-iteration method)	Xu, Davis, West
2013	Benchmarking studies with SOS, VLIDORT, MOM, & Monte Carlo	Davis et al.
2014	Integrated with a bio-optical model and multi-pixel optimization algorithm for ocean-color/aerosol combined retrieval	Xu, Dubovik, Zhai
2015 2016	Integrated with multi-pixel optimization algorithm for aerosol and land surface reflectance combined retrieval	Xu
2016	Formalism established for solving RT which accounts for unresolved random fluctuations of scattering particle density	Xu, Davis
2017	Combined with a line-by-line model and the double-k method to account for gaseous absorption	Xu

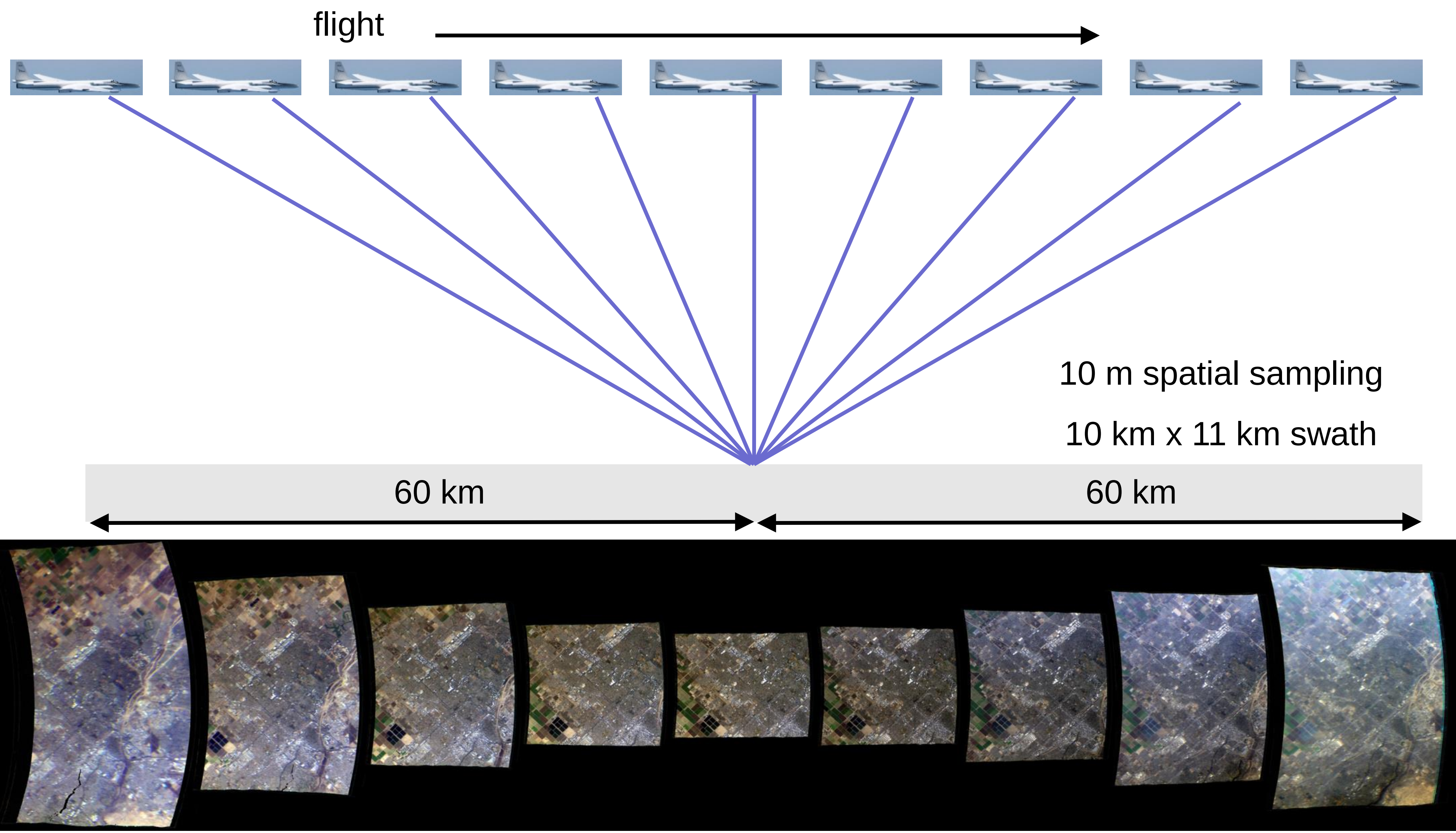
Thank you !

Airborne Multi-angle Spectro-Polarimetric Imager (AirMSPI)

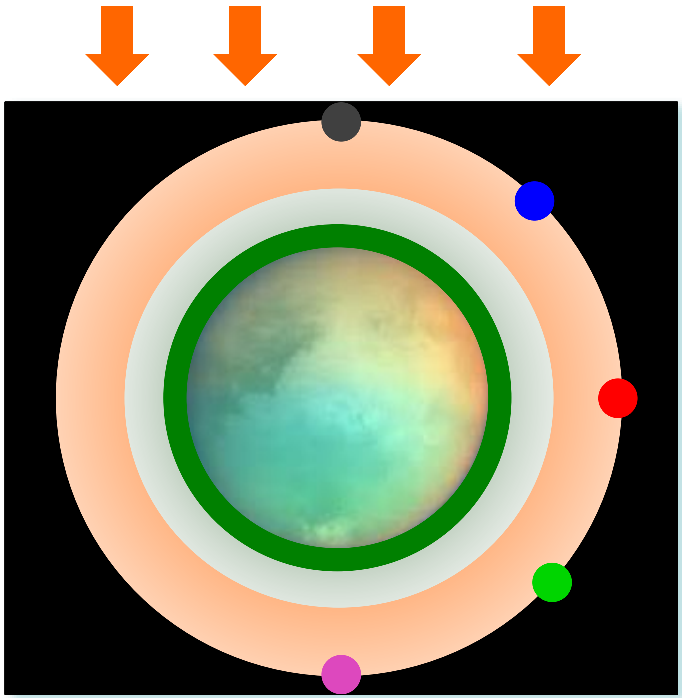


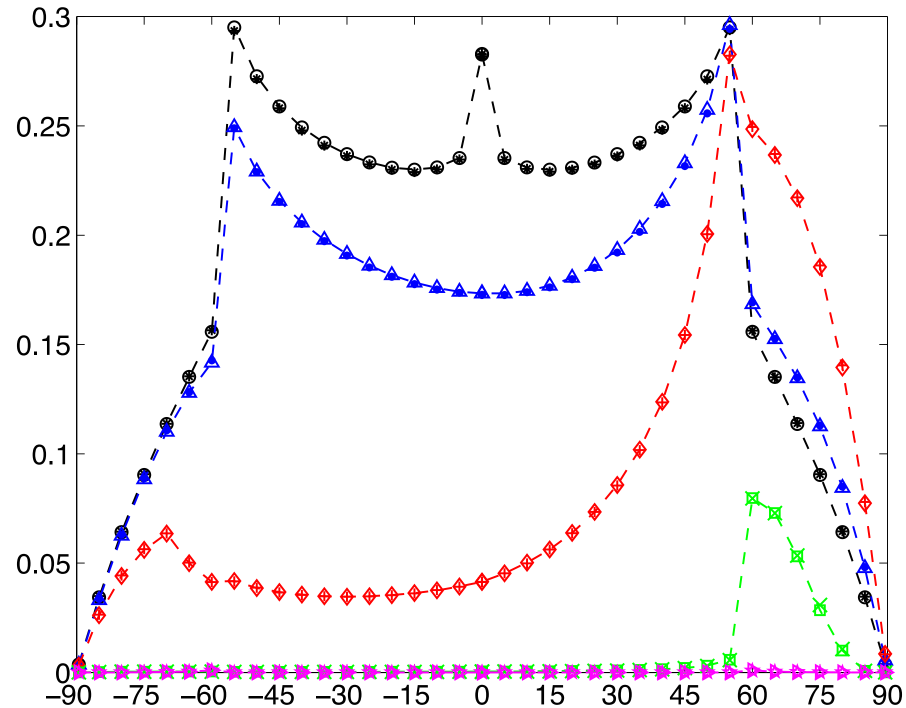
Spectral bands	355, 380, 445, 470*, 555, 660*, 865*, 935 nm (*polarized)
Platform	Flying on NASA ER-2 since 2010
Flight altitude	20 km
Multiangle viewing	Between $\pm 67^{\circ}$ using single-axis gimbal

AirMSPI step and stare



Verification by Monte Carlo (I - intensity)





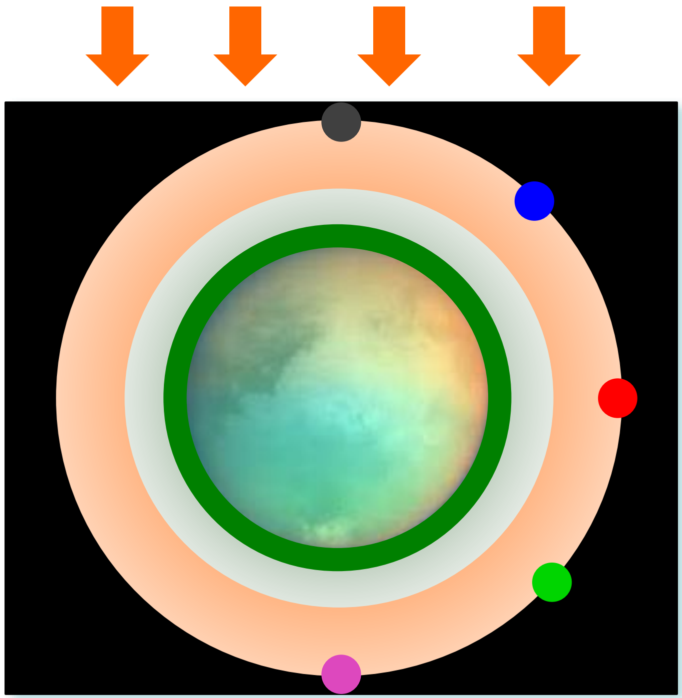
$\Theta = 0°$: MarCh | $\Theta = 45°$: MarCh | $\Theta = 90°$: MarCh | $\Theta = 135°$: MarCh | $\Theta = 180°$: MarCh

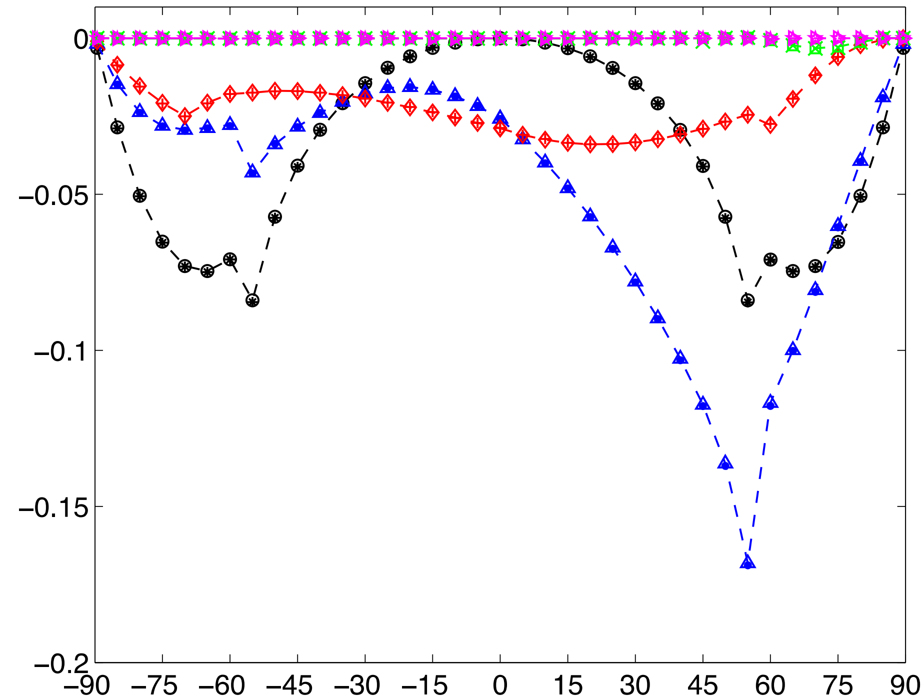
$\Theta = 0°$: Monte Carlo | $\Theta = 45°$: Monte Carlo | $\Theta = 90°$: Monte Carlo | $\Theta = 135°$: Monte Carlo | $\Theta = 180°$: Monte Carlo

Construction of scattering matrix: $P_{11}(a) = f_a P_{HG}(g_a, a) + (1 - f_a) P_{HG}(g_b, a)$ $\quad \frac{P_{ij}(a)}{P_{11}(a)} = f \left(\frac{P_{ij,\ \mathrm{Rayleigh}}(a)}{P_{11,\ \mathrm{Rayleigh}}(a)} \right)$

Braak et al. JQSRT Vol 69, 2001

Verification by Monte Carlo (Q)





Θ = 0°: MarCh
Θ = 45°: MarCh
Θ = 90°: MarCh
Θ = 135°: MarCh
Θ = 180°: MarCh

Θ = 0°: Monte Carlo
Θ = 45°: Monte Carlo
Θ = 90°: Monte Carlo
Θ = 135°: Monte Carlo
Θ = 180°: Monte Carlo

Xu et al. JQSRT Vol. 117, 2012

Time for Markov chain computation

Table. Execution time for adding/doubling and Markov chain methods*

Method	Rayleigh Atmosphere**	Inhomogeneous Venus Atmosphere
Adding/doubling	0.8 seconds	370 seconds
Markov chain	1.2 seconds	18 seconds

** Optical thickness: 84.

Why Markov chain is faster for inhomogeneous atmosphere?

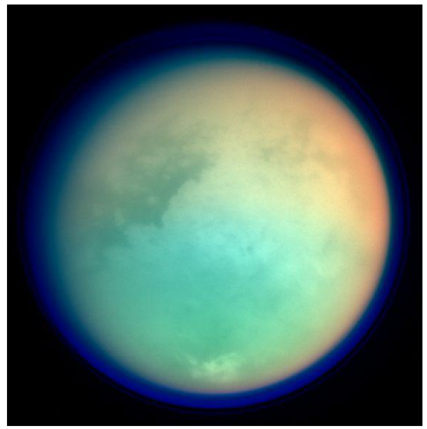
• Initial layer optical thickness can be as large as 0.01~0.05 as single scattering dominates

• In a chain, 10-20 layers are handled at one time adding/doubling method deals with 2 layers at one time

• "Adding" in doubling method means layer-to-layer adding while "adding" in Markov chain means chain-to-chain adding (here one "chain" means a sub-layer group)

* L. Esposito, Astrophys. J 233, 661 (1979).

Venus



Titan

In the time comparison table cited in my presentation, adding/doubling method is still faster than Markov chain method for homogeneous atmosphere. For plane-parallel atmosphere, it is slower. Nowadays, there should be some new techniques used in adding/doubling method to speed it up. But in principal, adding method is adding one more layer each time while Larry's adding concept is chain-to-chain adding. It is adding a chain (or a group of layers) at one time. So it should still retain some advantage in speed against the standard adding method.

Nevertheless, time comparison is still very tricky since (1) coding skills as well as (2) approximation made in different codes can influence the performance of speed. (3) One speed-up technique implemented in one RT method might not be used in another. Or (4) one speed-up technique implemented in one RT method is not be applicable to another. I did some speed comparison between Markov chain and VLIDORT but didn't show in my slides that Markov chain wins because of above reasons. So I directly cited Larry's table as the start of our Markov chain story. I should remind the audience that these comparison was long time ago and may need a revisit in a careful way.

Markov chain for radiation transfer

1: Status of photon: (n, i)

n: the layer number where photon stays

i: the direction it is going (θ_i)

2: Transition matrix $Q_{(n',j')|(n,i)}$

Transition probability of a photon from one intermediate status (n, i) to another (n', j) for next order of scattering;

3: Absorbing matrix $R_{(e)|(n',j)}$

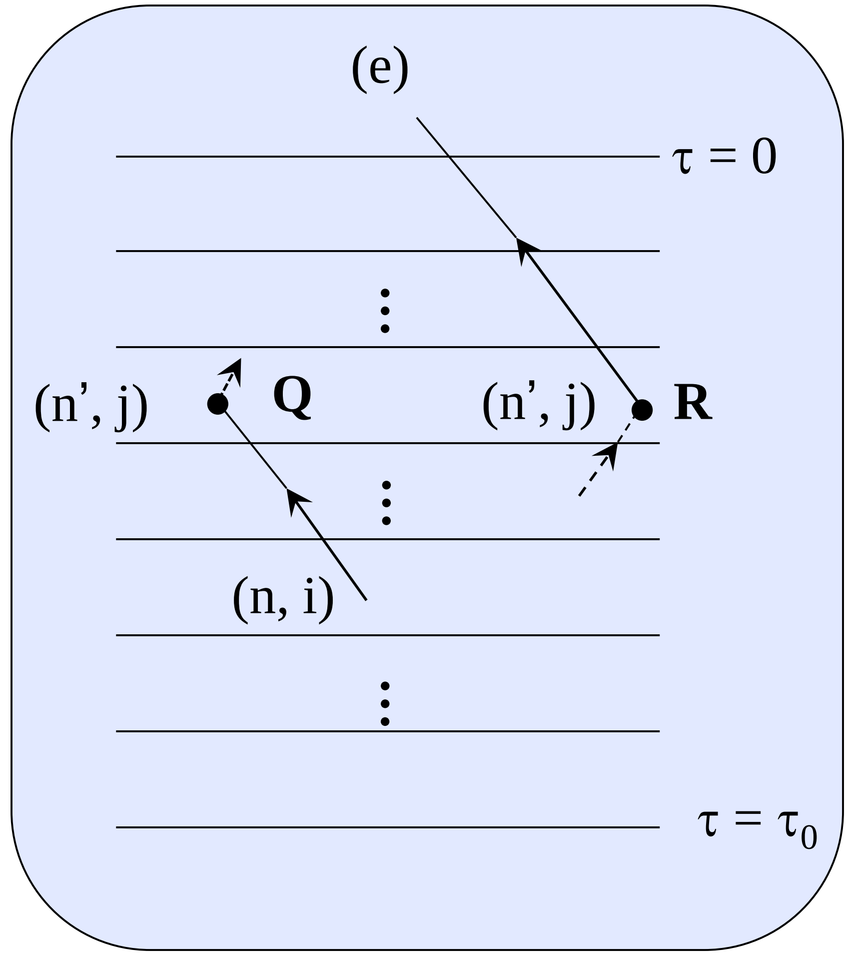
Probability of a photon escaping from intermediate status (n', j) to be out of the atmosphere in direction e (θ_e);

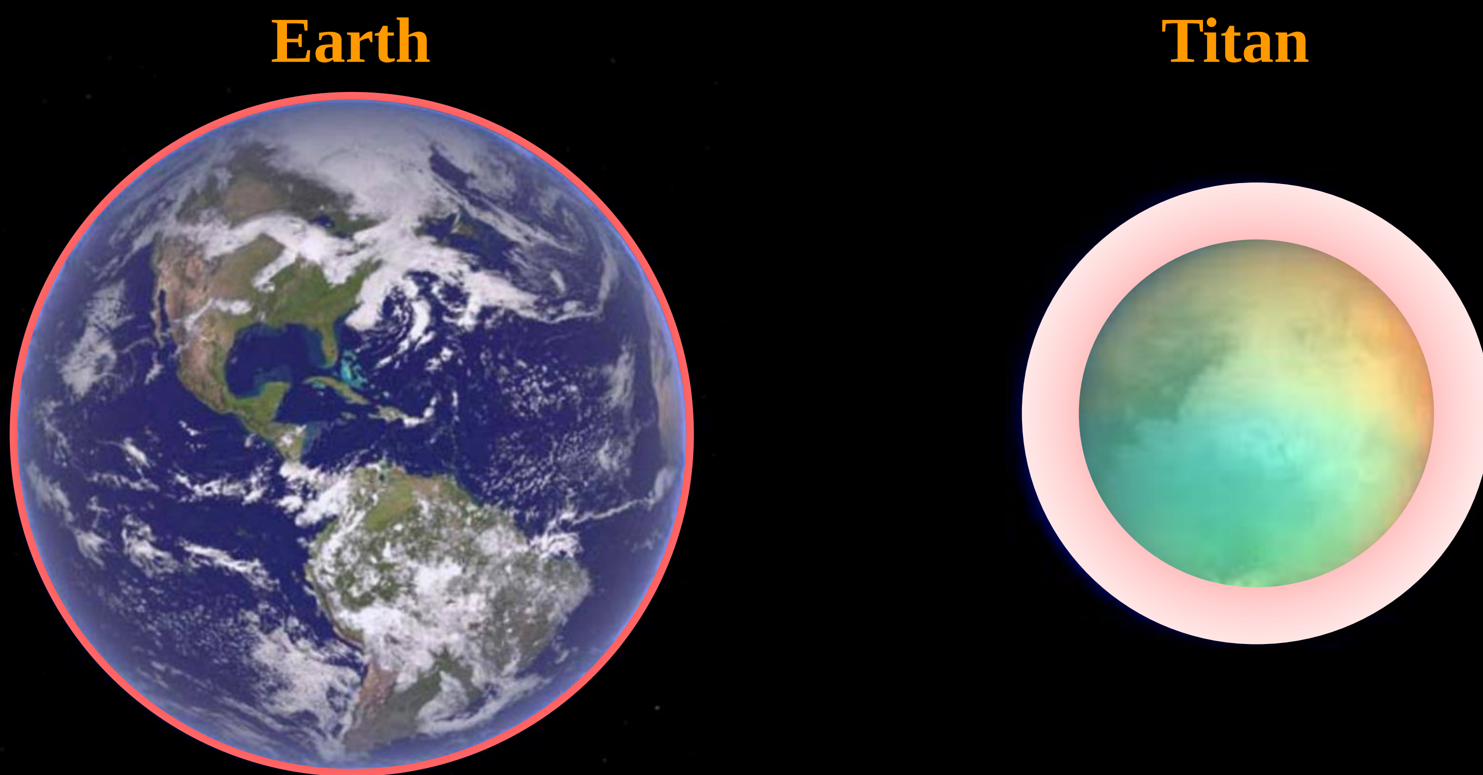
4: Initial probability distribution I_0

Photon's initial distribution in status (n_0, i_0)

5: Solution vector I:

The photon's probability to be emergent in direction (θ_e)





Parameter	Earth	Titan	Parameter	Earth	Titan
Surface Radius (km)	**6371**	**2575**	Surface Pressure (bar)	~1	1.5
Haze thickness (km)	**<20**	**~500**	Surface Temperature (k)	~300	~93
Surface gravity (cm s^{-2})	978	135	Density at the surface	1	~5
Solar Flux	1	0.011	Composition	~78% N_2 ~21% O_2	~95-98.5% N_2 1.4-4% CH_4